THE ROYAL SOCIETY

*Published online* 11 June 2003

# Abstraction and context in concept representation

**James A. Hampton**

*Department of Psychology, City University, Northampton Square, London EC1V 0HB, UK (hampton@city.ac.uk)*

This paper develops the notion of abstraction in the context of the psychology of concepts, and discusses its relation to context dependence in knowledge representation. Three general approaches to modelling conceptual knowledge from the domain of cognitive psychology are discussed, which serve to illustrate a theoretical dimension of increasing levels of abstraction.

**Keywords:** concepts; abstraction; prototype; exemplar; schema; instantiation

## 1. INTRODUCTION

Information is represented in the mind at widely different levels of abstraction. Consider how you might store information about a particular scene in a film, about the layout of the Paris Metro system or about Newton's first law of motion. If you are a movie fan, then your memory for the moment at the end of *Casablanca,* where Humphrey Bogart and Claude Rains are left on the airport runway, will be full of detail; the graininess of the film, the expressions of faces and voices are likely to be stored and recalled, even if the memories are inaccurate or distorted. Your knowledge of the Paris Metro will be a broader set of information—including some basic geography of Paris, notions of how distance can be translated into travel time, and specific details of individual lines or metro stations that you have passed through. When considering an understanding of Newton's laws of motion, we have a type of knowledge that may involve an ability to verbally express a formula ('a body in motion will continue to move in a straight line with constant velocity unless acted upon by some force'), together with an ability to use the law to make predictions and to understand the operation of a wide variety of physical systems, from cars braking to satellites orbiting the planet.

These three kinds of stored information represent different levels of abstraction in memory. We even have different ways of describing the information in our language—using terms such as memory, knowledge and understanding to refer to information varying from the specific to the abstract. When we think we have 'explained' or 'understood' something, we often mean that some specific event or situation is to be seen as an instantiation of some more abstract principle or notion.

One of the key notions underlying the concept of abstraction is the notion of *context independence*. Whereas the scene in *Casablanca* can only be understood in the context of that particular film, the laws of motion can be applied in almost any context. My aim, therefore, is to develop the notion of abstraction in more detail, and to discuss its relation to context dependence in knowledge representation. I will do this by discussing three general approaches to modelling conceptual knowledge from the domain of cognitive psychology, which themselves can be mapped onto a dimension of increasing levels of abstraction.

## 2. ABSTRACTION

The notion of abstraction is itself very abstract, and the term 'abstract' probably has many senses. For example, we distinguish concrete from abstract nouns in terms of whether they refer to physical objects or non-physical notions. Abstract concepts such as 'loyalty' or 'dimension' may refer to physical situations but they are unrelated to any direct sensory experience. Abstract art, however, is painting or sculpture that does not purport to represent the world in a recognizable form, but in which the sensory qualities of the paint or object are *primary*—a Rothko painting is 'just' patches of colour, a Brancusi sculpture just smooth shiny curves. Thus, we must be careful not to assume that there is just one notion of abstraction.

Let us begin with a simple example of abstraction, and then build up to more abstract levels. At its most basic, the process of abstraction may be considered as a form of *generalization* across time and place. Suppose that we set up a classical discrimination learning experiment in which a pigeon receives a food reward in the presence of one class of stimuli, and does not in the presence of a different class. For example, food may be available if a red key is pecked, but will be withdrawn if a yellow key is pecked. If we train the pigeon with red and yellow keys, we can then test its acquired 'knowledge' through a generalization test. How likely is the pigeon to peck keys of different shades of colour ranging between red and yellow? Typically, the level of response will show a generalization gradient—the pigeon will peck colours that have never been encountered before, to the extent that they are close or similar to the ones that were rewarded.

This form of learning employs abstraction because an experience is identified through linking it to some generic representation of similar experiences in the past (Millikan 1984). Our memories act to recognize the object or situation as a familiar one, and we respond accordingly.

One contribution of 16 to a Theme Issue 'The abstraction paths: from experience to concept'.

The abstraction involves storing three types of information:

(i) information about which dimensions of the situation are relevant (for example the colour rather than the shape or size of a key);
(ii) information about which values on which dimensions reliably predict how we should act (for example the red and yellow values on the colour dimension);
(iii) information about the range of variability of the predictive values (so if a range of reds is rewarded, a more 'abstract' representation of the stimulus class is created than if just a single value is rewarded, and there will be greater generalization of responding).

Crucially, abstraction also involves *not* storing anything else, or at least separating out this action-relevant information from the rest. So the representation will be more abstract the greater the degree to which *only* the important or relevant information is stored, and all else is discarded. Abstraction thus provides for rapid and easy processing of information—we are not distracted by irrelevant variation along other dimensions. By contrast, the down side of abstraction is that we fail to notice or record details, which may prove to be relevant should the task or situation change.

All abstraction therefore involves selectively discarding some of the information presented. So, for example, we may form an abstract representation of the concept TRIANGLE through selecting out the common elements (closed plane figure with just three straight sides and three corners) and ignoring details of the angles or sizes of individual triangles that we may have encountered or could imagine.

Along with this selection of relevant dimensions comes the construction of a *type* representation—a representation not of an individual or particular object, but of a class of possible objects. Individual objects may be encoded in memory with more or less attention to detail, and so give rise to memories that are more concrete or more abstract. Remembering that the bank robbers escaped in some kind of car will be a more abstract recollection than remembering that they drove off in a rusty red station wagon with a broken tail-light. However, the formation of a *type* is a different form of abstraction. Individual objects are *tokens* or instances of particular categories of thing. To represent that category as a type is to assume that there is something that such tokens have in common. We then set up a type representation that is more than just a collection of remembered exemplars.

The process of abstraction of a type involves us in making an ontological commitment that the collection of individuals we have encountered constitutes some form of coherent class. To illustrate this, let us consider a medical example. A physician may start encountering several patients with similar symptoms—headaches, night sweats and loss of appetite. Initially she may class them together purely on the basis of these symptoms. Such a category is convenient, because it can be given a name, and can be used as the basis for organizing case records. At this point, the physician may then decide to adopt a hypothesis that these cases have some common cause—there is, perhaps, some infectious agent, or some toxicity in the environment that is producing the cluster of similar cases. To form such a hypothesis is to create an abstract type representation for the category. Although behaviourists often speak of discrimination learning as a form of 'concept learning', one could argue that it is only with the construction of types that true concept formation occurs.

The final, and most abstract, stage in the development of abstract representations of the world is to construct a *type hierarchy*. Our conceptual knowledge is organized around an ontological framework (Keil 1979). We divide the world into domains of people, biological kinds, artefact kinds, individuals, qualities, events and so forth. In each domain, there is a form of template that dictates what relevant kinds of information should be stored. In the case of people, we may store their age, behaviour or beliefs. In the case of a car, age and behaviour are relevant but beliefs are not. To understand what questions it is relevant to ask about some new class, and what questions would just make no sense, we must abstract a framework in which there are types of types. One common assumption, used to good effect in object-oriented programming languages is that concept types are organized in a hierarchy (Keil 1979).[1]

## 3. THE ORIGINS OF ABSTRACTION

Empiricist and rationalist traditions in philosophy agree that the mind can form both specific and abstract representations. We are, at the same time, capable of appreciating the rich flavour of a particular sauce, and speculating about the status of ethical laws in human society. The argument between the traditions is about the origins of these representations and whether abstract structures are a necessary precursor to empirical learning.

This argument is also endemic in psychological theories of cognition. It has been argued that all knowledge must be grounded in some form of sensory experience, and that theories of learning are therefore crucial to understanding how knowledge comes about. By contrast, one can argue that without some form of abstract structure to give meaning to experience, learning would be random and impossible. Piaget was one of those who understood this problem most clearly (Piaget & Inhelder 1969). In his theory of development, the child begins with rudimentary structures—schemas—that allow action to be tied to sensory input (for example the sight of a face leads to tracking of the eyes). These schemas then develop and change by adapting themselves to the structure of experience, thus allowing richer meaning to be given to new inputs.

Psychological models of concept representation differ in the emphasis that they place on bottom-up learning processes as opposed to top-down interpretative processes. In the next section, I turn to the consideration of two broad classes of model that can be differentiated on this basis.

## 4. MODELS OF CONCEPTS

The two classes of model considered both use similarity as the basis for forming categories. Exemplar models do so in terms of stored individuals, whereas Prototype models do so by abstracting a single representation of the class.

### (a) *Exemplar models*

Exemplar models of concept representation are located at the 'specific' end of the abstraction dimension. There are two main lines of research employing these models. The first line is research into 'non-analytic' cognition developed by Lee Brooks and Bruce Whittlesea, and the second line involves constructing models of category learning based on representing exemplars in similarity spaces.

#### (i) *Non-analytic cognition*

Brooks (1978) proposed the radical thesis that much human behaviour that we assume to reflect conceptual or analytic thinking is in fact based on more specific memory-like processes. For example, we commonly find typical members of categories (for example cars as vehicles) easier to process than atypical members (such as hot-air balloons as vehicles). Rosch (1975) proposed that this is because of the similarity of different category members to some stored prototype for the category (see also Hampton 1979). However, it is equally possible that each time we encounter a vehicle we store some quite specific memory of it, and that the speed of processing typical category members reflects the relative frequency with which they are encountered, rather than anything that would require there to be a generic abstract representation of the category prototype. Research on this question (Barsalou 1985; Hampton 1997*a*) has shown that speed of categorization is influenced by both frequency of items in the world and their similarity to a prototype. Hampton (1997*a*) found that when the task of categorizing words was made very quick and easy, by having only unrelated false items as fillers, then frequency was the primary predictor of categorization speed. However, when the task required more cognitive effort to discriminate members from related non-members, similarity to prototype was more influential.

The non-analytic cognition tradition has provided many experiments demonstrating the strong effects of prior processing episodes on subsequent use of abstract knowledge. For example, even when the task is to classify stimuli on the basis of a simple rule, performance is affected by prior exposure to the stimuli (Allen & Brooks 1991). For a review of many of these demonstrations, see Shanks (1995).

Brooks *et al.* (1991) conducted a study of doctors' diagnoses of common skin diseases. The participants in the study first worked their way through named slides of skin disorders of various categories, and rated how typical they were of their disease category. Later, they were given a new set of slides to categorize according to the diagnostic categories. Brooks *et al.* (1991) found that the preferred diagnosis was consistently influenced by the similarity of a novel case to the recently viewed cases of that category. Even though the experts had been trained in the application of abstract rules for classification, they were unable to avoid using similarity to recently experienced exemplars to help them categorize. Brooks concludes that perhaps much of the behaviour that we assume is based on abstract representations is actually exemplar based.

This approach is the most clearly *anti-abstractionist*. It makes no assumptions about representation, other than that a stimulus may evoke the memory of an earlier one together with its processing history. All that is required for a model of concept representation is some way in which a novel stimulus or experience can evoke the right kind of earlier episode—a principle of 'reminding'. Such a system must presumably be mediated by a representational medium of some form. Reminders are typically triggered by salient perceptual features of a stimulus rather than abstract or functional features (Ross *et al.* 1990), so it must be assumed that there has been some selective storage of information. However, the force of the demonstrations is that this selection is probably automatic and driven by basic perceptual attention processes, and is not easily modified by the imposition of strategies for the performance of the task.

One clever way to determine how much abstraction is involved in a task is to use an implicit learning paradigm. Barsalou & Ross (1986) presented participants with lists of words to read. At a later stage, they were asked to estimate the frequency with which particular classes of word had appeared. Frequencies for categories based on simple perceptual features (e.g. red objects) were not estimated better than chance—participants were unable to say reliably whether there had been one, two, three or four red objects in the list. However, frequencies for common superordinate categories such as Birds, Fruits or Vehicles were well estimated. A plausible interpretation of this result is that there is some automatic encoding of a word's superordinate category when it is read and understood. If so, then this would be evidence for a degree of automatic abstraction occurring during the processing of the word's meaning.

#### (ii) *Generalized context model*

Perhaps the best known and most influential exemplar model for concept representation is the GCM developed by Robert Nosofsky (1988). The GCM is a model of classification learning and has largely been tested through studies in which participants have to learn to classify a set of simple visual shapes into two categories. Measures are taken of the rate of learning of individual stimuli, and a test is made of generalization of the learning to a transfer set of stimuli that were not seen in training.

The model assumes a very simple system for representing a category, involving a minimum of abstraction. Individual exemplars are stored together with their category labels, and the subsequent classification of a new stimulus is based on the overall similarity of the stimulus to the stored exemplars of each category in memory.

In more detail, the GCM assumes that each exemplar is analysed along a set of dimensions (experimental tests of the model typically use visual stimuli that can be fully defined by their position on just two independent dimensions such as size and orientation). By plotting a space with the dimensions as axes, it is then possible to store an exemplar and its category by placing the label of that category at the point in the space corresponding to the exemplar's dimensional coordinates. Once the space has become populated with exemplars, the categorization of a novel stimulus can be made by comparing the sum of similarities to members of category A with the sum of similarities to members of category B. The probability of categorizing the stimulus as an A is predicted directly through a formula comparing these two summed similarities.

Formally, the model uses several mathematical devices, details of which can be found in Nosofsky (1988) or Lamberts (1997). To fit the data from any experiment, it is also necessary to use the data to estimate several parameters for the model. These include the following: the relative importance of each dimension; the rate at which similarity to a stored exemplar drops off as a function of distance in the space; and the degree to which the probability of choosing a category is maximized (always choosing the more likely category) or follows probability matching (responding with a probability that matches the likelihood of the category being correct).

The GCM has had considerable success in giving precise quantitative fits to data from a range of different experimental paradigms, including recognition accuracy for individual exemplars, and speed of categorization (Lamberts 1995). More recently, the generality of the model has been challenged by J. David Smith, who has demonstrated a range of situations in which performance indicates that people may abstract a prototype rather than store individual exemplars (Smith & Minda 1998). Broadly speaking, where a category has many exemplars, and the differentiation between categories is relatively easy (large distances between categories relative to variance within), then there tends to be more evidence for the formation of a category based around a prototype (a single point in the stimulus space) rather than exemplars.

Exemplar models employ a degree of abstraction, because only the information about position in the similarity space is encoded for any stimulus, all other information is lost. Where there are additional dimensions of variation that are not relevant to the classification, it can be shown that attention shifts away from these dimensions and they lose any influence in determining performance (Goldstone 1994). A model that successfully captures the learning of dimensional weights during classification learning is the ALCOVE model developed by John Kruschke (1992). Goldstone has demonstrated that this form of learning can affect visual discrimination abilities—people retune their perceptual processing of the stimuli to attend to the dimensions that are more relevant.

To what extent is it possible to abstract out from a perceptual array only those dimensions that are relevant to a discrimination? An answer to this question can be found in the work of Garner (1978) who developed the distinction between separable and integral dimensions. If a pair of dimensions are separable (such as shape and colour) then it is possible to attend to only one dimension, and variability on the other dimension will not interfere with processing. Furthermore, similarity between stimuli will be the inverse of an additive function of distance on each dimension. Where a pair of dimensions are integral (such as hue and saturation in colour) then variability in one dimension will always interfere with judgements on the other, and similarity between stimuli is best captured by using a Euclidean function (similarity is inversely related to the square root of the sum of the squared distance on each dimension).

#### (iii) *Exemplars and natural categories*

There have been relatively few attempts to generalize the GCM beyond the realm of tightly controlled laboratory experiments using simple visual stimuli. A notable exception is the work of Gert Storms and his colleagues. For example, Storms *et al.* (2000) compared two models of category internal structure. For a prototype model (see next section), they predicted that typical category members would be those with the most features in common with the category prototype (Hampton 1979). For an exemplar model, they predicted that typical category members would be those that shared most features with the top few most frequently generated members. Their results showed that for a range of measures of category structure (speed of categorization, rated typicality and so forth), the exemplar measure was as good as, or in some cases better, than the prototype measure. Optimum fit was obtained when combining similarity to approximately 10 exemplars.

Another demonstration of exemplar effects was a study by Storms *et al.* (2001) in which they examined how people classify edible plant objects into fruits and vegetables (or their equivalent in Flemish). They obtained a range of exotic produce that was unknown to their participants and obtained ratings of similarity to familiar fruits and vegetables, and judgements of which category they belonged to. They found that similarity to known exemplars was a better predictor of classification than was the degree of feature match with the category prototype.

### (b) *Prototype models*

Prototype models stand in the 'middle ground' between non-analytic exemplar storage and highly abstract 'theory-like' conceptual representations. A variety of representational assumptions can be adopted. For example, one can assume that stimuli are represented in a multidimensional similarity space (Rips *et al.* 1973) of the kind also used by exemplar models. Or one can assume a more powerful schema-based representational format (e.g. Smith & Osherson 1984; Hampton 1987, 1995*a*).

#### (i) *Spatial prototype models*

The key difference between prototype and exemplar models of classification learning is that prototype models involve abstraction over exemplars to represent the central tendency and the variability within the category. Rather than represent each exemplar as a point in the similarity space, the prototype model would represent the centre of each category cluster and define the category as a region in the space centred on that point. Individual exemplars are therefore merged into a single generic representation.

There are two effects of using a prototype representation. First, the likelihood of selecting a category will decrease in a smooth monotonic fashion as distance from the centre of the category increases. There will be no local maxima in the regions close to individual exemplars as occur in the exemplar model. Second, information about higher-order covariance of features across exemplars is lost. The bivariate distribution of exemplars across two dimensions of the space may show a correlation, but the prototype only represents the centre of the cluster and not its shape.

#### (ii) *Exemplars versus prototypes*

Prototype models do better at explaining learning when the stimuli are more complex, when there are more of

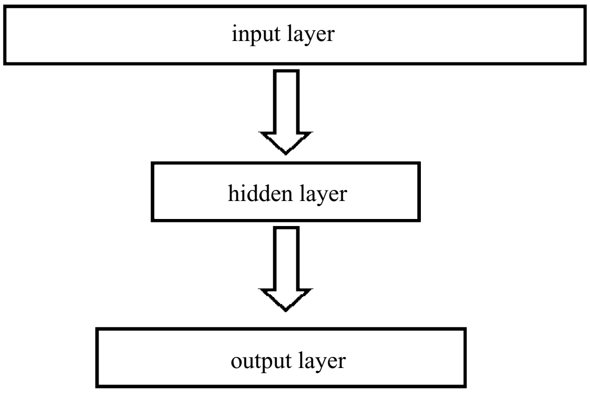


Figure 1. A simple feed-forward network for learning to categorize on the basis of input features. The more nodes allowed within the hidden layer, the more the network is able to learn the characteristics of individual exemplars. The more the number of hidden layer nodes is restricted, the more the network forms generalizations of the stimulus set. The degree of abstraction is thus related to the size of the hidden layer.

them, and when the category structure is well differentiated (Smith & Minda 1998).

A clear way to show how the difference between the models can be understood as one of degree of abstraction is by implementing the models within the same structure of a neural net (see figure 1). Barsalou (1990) argued persuasively that exemplar and prototype representations differ only in the degree to which individual exemplar information is retained (exemplar model) or discarded (prototype model). If one considers the simple feed-forward neural network in figure 1, the input layer encodes the particular set of features of different exemplars, the network carries this pattern through on the basis of the weights of each connection into a hidden layer, and then into an output layer where the different nodes each correspond to a different category. It is easy to show that, in such a model, the degree to which it behaves like an exemplar or like a prototype model is simply a function of how narrow the hidden layer is made to be. If the hidden layer has very few nodes, then the model necessarily has to throw away individual information about particular exemplars. There are not enough hidden layer nodes for each possible pattern of input to generate a distinctive pattern in the hidden layer, and so abstraction is forced on the representation. By contrast, if there are enough hidden layer nodes, each exemplar can be mapped onto a particular pattern in the hidden layer, and the category response then mapped onto that.

This aspect of the way in which neural networks learn is well understood, and when training them it is necessary to choose an appropriate number of hidden layer nodes to achieve the desired level of learning. Too many nodes leads to rapid learning on the training set but very poor generalization to a transfer set. Too few may lead to slow and incomplete learning of the training set, with some exceptional items never properly learned, but generalization to the transfer set will be improved. The hidden layer acts in statistical terms to reduce the dimensionality of the variance in the stimulus array, in a way that will be maximally useful for predicting the correct categorization.

(iii) *The effects of increasing expertise*

As learning continues, so more exemplar information is likely to become stored. Smith & Minda (1998) found that prototype models were a better fit early in learning, whereas exemplar models were better at a later stage. It is easier to obtain a general idea of two broad classes first, but achieving a more accurate level of performance requires the learning of individuals. This notion, that exemplar representations come with increasing expertise, has some support from studies of expertise in decision-making. In complex domains where knowledge and its application are uncertain, experts have been shown to have a greater store of relevant experience than well-trained novices, and to rely to a greater extent on stored exemplars (using a process known as 'case-based reasoning') in reaching decisions. Rosch *et al.* (1976) also argued that the perception of a class as having high commonality—associated with their notion of a 'basic level class'—can change with expertise in the same manner. The novice student of art history may see paintings by Vermeer as a single recognizable category, based on a prototype of the typical qualities of light and subject matter. To the expert, all the known individual paintings by the master are represented as familiar individual exemplars, so that the prototype representation no longer serves any great purpose.

(iv) *Schema-based prototypes*

Prototype models were originally developed for classification learning of visual stimuli (Posner & Keele 1968), but they were quickly adapted as models of common everyday concepts like Chair or Fruit (Rosch 1975; Hampton 1979). Of course the variety of Fruits require that the dimensionality of a similarity space would be very large, and in fact there is good evidence that the assumptions required for mapping similarities into multidimensional space are often violated by more complex natural concepts (Tversky 1977). It is therefore best to abandon the assumption of a similarity *space* as the medium for representation. In its place, concepts such as these are commonly represented using Minsky's idea of a frame with slots (= features or dimensions) and fillers (= values or ranges) (Minsky 1975).

For example, figure 2 shows part of a possible frame representation for the concept of APPLE. Slots represent variables that will differentiate different concepts within a particular domain. Thus, for fruits and vegetables, they would represent shape, colour, taste, origin and so forth. For vehicles, they might represent form of locomotion (wheels or wings), shape, use (passengers or freight) and location (air, land or sea). The particular set of slots for a domain will reflect the level of abstraction described above as an ontological framework. Thus depending on the location of a concept within the ontological hierarchy it may be expected to use a particular range of slots (although the frame representation is likely to be very flexible and expandable; Barsalou & Hale 1993).

These schema representations are *prototype* models, in that a single representation is used to store the central tendency and range of variation in the category across relevant dimensions, but no exemplar information is stored. They also have the important characteristic of prototype models, that although the most typical exemplars will be

| slot | values |
|---|---|
| shape | spherical with stalk |
| origin | trees |
| size | range from 4 to 15 cm |
| colour | distribution across red, green, brown, yellow |
| function | eaten |
| sub function | eaten on own, in pies, in sauce |
| taste | sweet, acidic |

Figure 2. A (partial) frame representation for the concept APPLE.

clearly described, the schema itself will not determine precisely the boundaries of the category. This vagueness in the placing of category boundaries differentiates schematic prototypes from more precise rule-based or well-defined concept representations.

Barsalou & Hale (1993) also describe how more powerful representations can be constructed around frame-based schemas by introducing constraints and meta-conceptual principles defined across the slots in the frame. Abstract knowledge—of the kind sometimes referred to as 'naive theories'—can be incorporated into the representation in this way. For example, within the domain of biological kinds there are constraints that operate between size, habitat, shape and means of locomotion. Small creatures may have wings and fly in the air. Creatures of any size may swim through the water by moving a tail. There is good evidence that we are aware of, and represent many of these higher-order relations between different features within a representation (Murphy & Medin 1985), although it is unclear to what extent we may be able to articulate such abstract knowledge in an explicit form.

Schematic prototypes therefore typically involve:

(i) a frame with slots and values;
(ii) constraints operating between dimensions that reflect the broader principles of world knowledge;
(iii) categorization rules that weight different dimensions in relation to their 'centrality' in the structure (as determined by the constraints).

The models tend to be imprecise about how a stimulus might be encoded, and about how categorization is decided. Rips (1989) suggested that we categorize something by finding the conceptual schema that provides the 'best explanation' of the observable features of the object. There is also evidence (Ahn 1998) that, given a schema in which some features are represented as the cause of others, people will give greater weight to the causally active features and less to their effects when categorizing. However, much needs to be done to discover how widespread such effects may be in the categories that we use in everyday reasoning. Some research (Malt 1990; Hampton 1995*b*) suggests that people are often quite indiscriminate in the way that they attribute weight to different kinds of information when deciding how to categorize an object.

## 5. CONTEXTUAL EFFECTS IN CONCEPT REPRESENTATION

Having described two basic classes of model on the basis of different levels of abstraction, I now turn to the question of context dependence. Contextual effects in conceptual tasks are *prima facie* evidence for the lack of abstraction. We have already discussed demonstrations by Brooks of the effects of the context of recent experience on performing categorization tasks. However, there are other forms of context manipulation that can have equally strong effects.

Exemplar and simple non-schematic prototype models are similarly affected by specific context effects. In particular, the order in which exemplars are presented early on in learning and the separability of the relevant dimensions may have effects on how quickly the category is learned. However, there are much more interesting context effects to be found with schema-based concept representations. I argue that such effects frequently occur through a process of *instantiation*, which can be thought of as the filling out of abstract representations with more specific features to help the concept to fit into the current context.

### (a) *Context-dependent properties*

One powerful effect of context is the addition of *new dimensions* to a representation. Barsalou (1982) used the example of a basketball. Participants had to verify two kinds of sentence. Some were constructed using highly associated properties as predicates—for example, 'a basketball is round'. Others were constructed using weakly associated properties that were still nonetheless true, such as 'a basketball floats'. Without any prior context, the former were more rapidly verified than the latter. However, if in the context of a prior sentence, such as 'Harry threw the basketball into the pool', the two sentences became equally fast to understand.

One way to interpret this result is to suppose that in the context of understanding the prior sentence about a swimming pool, the attribute 'floats' is added to the representation of basketballs. A particular individual basketball is represented in working memory, and its roundness is part of that representation regardless of the prior context. Its ability to float is added to the representation to provide coherence to the storyline, and allow a successful *simulation* in working memory of the situation described. Barsalou (1999) has since greatly developed the notion of mental simulation in regard to concepts, and argues that our basic conceptual knowledge consists of a set of abilities to mentally simulate objects in situations.

The context of placing a word in a sentence may also lead to the *negating* of common attributes. Consider the sentence 'The family watched hungrily as the cook took the bird out of the oven'. Clearly attributes of birds such as flying, having feathers and singing in trees will be absent from the working memory representation of the word in this context, although it is not known whether they are initially present and then deleted, or whether it is possible for the modified representation to be constructed directly.

### (b) *Instantiation as the reverse of abstraction*

It is possible to argue that most context effects involve the *instantiation* of an abstract concept by reference to plausibility in the context. For example, Roth & Shoben (1983) presented sentences such as 'the trucker sipped the beverage' and 'the bird crossed the farmyard'. They showed that in such circumstances, there was considerable priming of words that were plausible instantiations of the abstract category term (for example coffee, or chicken). There is a clear link here with Rosch and Mervis' notion of the basic level of concepts. The basic level in any conceptual hierarchy picks out the concept terms that are most general, but still easily imagined, in terms of a visual image. For example, the following are basic terms: car, chair, penguin or banana. Basic level terms are the first nouns that children typically learn, and they have been shown to have a wide range of processing advantages over more superordinate terms such as vehicle, furniture or fruit.[2] Instantiation in context appears to involve the retrieval of a basic level concept that will plausibly replace the superordinate. We can easily imagine a cup of coffee in the trucker's hand, but not so easily an unspecified beverage. Interestingly, it is probably rare that instantiation would venture to be more specific than the basic level—we do not appear to fill in what colour the chicken was or whether the coffee was white or black when understanding the sentences.

A second example of context effects involving instantiation is some research by Barsalou (1987). He had the original idea of asking students to adopt different points of view when making judgements about the typicality of different members of a category. For example, they might have been asked to judge a list of vehicles for how typical they were from the point of view of a suburban housewife as opposed to the point of view of a farmer. In this case, the students were able to agree on a consistent notion of how typical the objects would seem to the target group, and the ranking of typicalities changed radically according to point of view. Indeed, when asked to make judgements from the point of view of a university professor, they were in reasonable agreement with the actual views of professors measured independently (although the professors were apparently less good at matching the views of students).

How to explain these effects? A shift in typicality gradients might indicate that a different prototype has been formed, but it is unclear how the students would be able to know what kind of prototype a farmer or housewife might have. A more probable account is to suppose that the students were relying on exemplar representations, using their knowledge of the likelihood of finding trucks, tractors, estate cars or sports cars in the environment of a farm or suburb.

A third example of instantiation processes comes from the author's own studies of conceptual combination using conjunctively defined conceptual categories (Hampton 1987, 1988, 1997*b*,*c*). If people are asked to list attributes of birds, and separately to list attributes of pets, they will produce a list of about 30 different common properties. In Hampton (1987), I investigated (among other conjunctions) which of these properties would then be considered true of the conjunction 'birds which are pets'. A striking result was that people generated attributes for the conjunction that had not been considered true of either class considered alone. Birds do not live in cages and nor do pets, but clearly pet birds do. People also claimed that pet birds can talk, although other subjects had judged it impossible for pets to talk. Once again, once the two superordinate categories were placed in context, particular basic level concepts were instantiated—in this case, concepts such as PARROT. The 'emergent features' of conceptual conjunctions were mostly of this kind—properties that were true of high probability instantiations of the conceptual conjunction.

An even stronger effect was found when negated conjunctions were studied (Hampton 1997*b*). For example, the concept 'dwellings which are not buildings' was instantiated (implicitly) as tents and caravans, and properties generated for the conjunction reflected the common features of these objects.

Sometimes it seems impossible to perform some conceptual tasks *without* instantiating abstract categories into more basic level ones. An example is the 'impossible object task' (Hampton 1997*c*). In this task, people are asked to imagine an object that does not exist, such as 'A piece of furniture which is also a fruit'.

They are encouraged to list properties of the object, to consider ways in which it differs from standard furniture or standard fruit, and to draw a picture to illustrate it if they wish.

The task is difficult, and only about half the participants typically find a reasonable resolution of the problem. Where successful, solutions appear to involve the following four elements:

(i) a search for instantiations of furniture and of fruit;
(ii) *alignment* of schemas;
(iii) identification of *conflicting elements*;
(iv) invention of modifications to reduce the conflict.

For example, they may choose to instantiate furniture as a chair, and then search for some fruit that could be fashioned into a chair. They may then look for something strong and large with a flat surface that could be sat upon, and come up with a pumpkin. At this point stage (ii) involves finding links between the parts and functions of one concept and those of the other. At various points in this process, the same slot needs to be filled with values that are incompatible. For example, chairs need to be long lasting, whereas pumpkins will start to rot. The final stage then involves a creative process of modifying one or other concept to ensure the resolution of the conflict—for example, people may declare that the chair will need replacing regularly but could be turned into a nice soup, or alternatively they may declare that genetic modification of the pumpkin has led to a fruit that remains hard and unripe for many years without a problem.

Most striking about these results is that people never create a new kind of furniture, or a new kind of fruit. There is almost always an instantiation of a familiar type, although it may not always be typical (depending on the particular constraints imposed by the other category).

## 6. CONCLUSIONS

In this paper, I have argued that abstraction within conceptual representations involves three levels:

(i) the selective storage of relevant information and discarding of irrelevant information;
(ii) the development of type rather than token representations;
(iii) the development of higher-order constraints within an ontological hierarchy.

Psychological models of concepts are distinguishable by the level of abstraction that they incorporate and, in particular, exemplar models operate at a much lower level of abstraction than prototype models. There is considerable evidence that people operate at a very specific level of information processing, even when faced with relatively abstract tasks. Furthermore, a range of context effects are best understood by supposing that when presented with broad superordinate terms, there is a preference for instantiating them at the basic level.

## ENDNOTES

[1]An additional form of abstraction is the creation of metarepresentations—representations of representations. This type of abstraction is very important in mathematics and several sciences, but I will not discuss it here.

[2]Bird is an exception here—it appears from the evidence that both 'bird' and 'penguin' are basic level terms—a complication that there is not space to explain here.

## REFERENCES


Ahn, W. K. 1998 Why are different features central for natural kinds and artifacts? The role of causal status in determining feature centrality. *Cognition* **69**, 135–178.

Allen, S. W. & Brooks, L. R. 1991 Specializing the operation of an explicit rule. *J. Exp. Psychol. Gen.* **120**, 3–19.

Barsalou, L. W. 1982 Context-independent and context-dependent information in concepts. *Mem. Cogn.* **10**, 82–93.

Barsalou, L. W. 1985 Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *J. Exp. Psychol. Learn. Mem. Cogn.* **11**, 629–654.

Barsalou, L. W. 1987 The instability of graded structure: implications for the nature of concepts. In *Concepts and conceptual development: ecological and intellectual factors in categorization* (ed. U. Neisser), pp. 101–140. Cambridge University Press.

Barsalou, L. W. 1990 On the indistinguishability of exemplar memory and abstraction in category representation. In *Advances in social cognition*, vol. III (ed. T. Skrull & R. S. Wyer), pp. 61–88. Hillsdale, NJ: Lawrence Erlbaum Associates.

Barsalou, L. W. 1999 Perceptual symbol systems. *Behav. Brain Sci.* **22**, 577–609.

Barsalou, L. W. & Hale, C. R. 1993 Components of conceptual representation: from feature lists to recursive frames. In *Categories and concepts: theoretical views and inductive data analysis* (ed. I. van Mechelen, J. A. Hampton, R. S. Michalski & P. Theuns), pp. 97–144. London: Academic Press.

Barsalou, L. W. & Ross, B. H. 1986 The roles of automatic and strategic processing in sensitivity to superordinate and property frequency. *J. Exp. Psychol. Learn. Mem. Cogn.* **12**, 116–134.

Brooks, L. R. 1978 Non-analytic concept formation and memory for instances. In *Cognition and categorization* (ed. E. R. Rosch & B. B. Lloyd), pp. 170–211. Hillsdale, NJ: Lawrence Erlbaum Associates.

Brooks, L. R., Norman, G. R. & Allen, S. W. 1991 Role of specific similarity in a medical diagnosis task. *J. Exp. Psychol. Gen.* **120**, 278–287.

Garner, W. R. 1978 Aspects of a stimulus: features, dimensions, and configurations. In *Cognition and categorization* (ed. E. Rosch & B. B. Lloyd), pp. 99–139. Hillsdale, NJ: Lawrence Erlbaum Associates.

Goldstone, R. L. 1994 The role of similarity in categorization: providing a groundwork. *Cognition* **52**, 125–157.

Hampton, J. A. 1979 Polymorphous concepts in semantic memory. *J. Verbal Learn. Verbal Behav.* **18**, 441–461.

Hampton, J. A. 1987 Inheritance of attributes in natural concept conjunctions. *Mem. Cogn.* **15**, 55–71.

Hampton, J. A. 1988 Overextension of conjunctive concepts: evidence for a unitary model of concept typicality and class inclusion. *J. Exp. Psychol. Learn. Mem. Cogn.* **14**, 12–32.

Hampton, J. A. 1995*a* Testing prototype theory of concepts. *J. Mem. Lang.* **34**, 686–708.

Hampton, J. A. 1995*b* Similarity-based categorization: the development of prototype theory. *Psychol. Belg.* **35**, 103–125.

Hampton, J. A. 1997*a* Associative and similarity-based processes in categorization decisions. *Mem. Cogn.* **25**, 625–640.

Hampton, J. A. 1997*b* Conceptual combination: conjunction and negation of natural concepts. *Mem. Cogn.* **25**, 888–909.

Hampton, J. A. 1997*c* Emergent attributes in conceptual combinations. In *Creative thought: an investigation of conceptual structures and processes* (ed. T. B. Ward, S. M. Smith & J. Viad), pp. 83–110. Washington, DC: American Psychological Association Press.

Keil, F. C. 1979 *Semantic and conceptual development: an ontological perspective*. Cambridge, MA: Harvard University Press.

Kruschke, J. K. 1992 ALCOVE: an exemplar-based connectionist model of category learning. *Psychol. Rev.* **99**, 22–44.

Lamberts, K. 1995 Categorization under time pressure. *J. Exp. Psychol. Gen.* **124**, 161–180.

Lamberts, K. 1997 Process models of categorization. In *Knowledge, concepts, and categories* (ed. K. Lamberts & D. Shanks), pp. 371–403. Hove: Psychology Press.

Malt, B. C. 1990 Features and beliefs in the mental representation of categories. *J. Mem. Lang.* **29**, 289–315.

Millikan, R. 1984 *Language, thought and other biological categories*. Cambridge, MA: MIT Press.

Minsky, M. 1975 A framework for representing knowledge. In *The psychology of computer vision* (ed. P. H. Winston), pp. 211–277. New York: McGraw-Hill.

Murphy, G. L. & Medin, D. L. 1985 The role of theories in conceptual coherence. *Psychol. Rev.* **92**, 289–316.

Nosofsky, R. M. 1988 Exemplar-based accounts of relations between classification, recognition, and typicality. *J. Exp. Psychol. Learn. Mem. Cogn.* **14**, 700–708.

Piaget, J. & Inhelder, B. 1969 *The psychology of the child*. New York: Basic Books.

Posner, M. I. & Keele, S. W. 1968 On the genesis of abstract ideas. *J. Exp. Psychol.* **77**, 353–363.

Rips, L. J. 1989 Similarity, typicality and categorization. In *Similarity and analogical reasoning* (ed. S. Vosniadou & A. Ortony), pp. 21–59. Cambridge University Press.

Rips, L. J., Shoben, E. J. & Smith, E. E. 1973 Semantic distance and the verification of semantic relations. *J. Verbal Learn. Verbal Behav.* **12**, 1–20.

Rosch, E. 1975 Cognitive representations of semantic categories. *J. Exp. Psychol. Gen.* **104**, 192–232.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M. & Boyes-Braem, P. 1976 Basic objects in natural categories. *Cogn. Psychol.* **8**, 382–439.

Ross, B. H., Perkins, S. J. & Tenpenny, P. L. 1990 Reminding-based category learning. *Cogn. Psychol.* **22**, 460–492.

Roth, E. M. & Shoben, E. J. 1983 The effect of context on the structure of categories. *Cogn. Psychol.* **15**, 346–378.

Shanks, D. R. 1995 Representation of categories and concepts in memory. In *Cognitive models of memory* (ed. M. A. Conway). Hove, UK: Psychology Press.

Smith, E. E. & Osherson, D. N. 1984 Conceptual combination with prototype concepts. *Cogn. Sci.* **8**, 337–361.

Smith, J. D. & Minda, J. P. 1998 Prototypes in the mist: the early epochs of category learning. *J. Exp. Psychol. Learn. Mem. Cogn.* **24**, 1411–1436.

Storms, G., De Boeck, P. & Ruts, W. 2000 Prototype and exemplar based information in natural language categories. *J. Mem. Lang.* **42**, 51–73.

Storms, G., De Boeck, P. & Ruts, W. 2001 Categorization of unknown stimuli in well-known natural concepts: a case study. *Psychon. Bull. Rev.* **8**, 377–384.

Tversky, A. 1977 Features of similarity. *Psychol. Rev.* **84**, 327–352.

## GLOSSARY

GCM: generalized context model